Infection, Genetics and Evolution 20 (2013) 118–123

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Short communication

Complete genome sequence and molecular phylogeny of a newfound hantavirus harbored by the Doucet's musk shrew (*Crocidura douceti*) in Guinea

Se Hun Gu [a], Violaine Nicolas [b], Aude Lalis [b], Nuankanya Sathirapongsasuti [a], Richard Yanagihara [a,*]

[a] Departments of Pediatrics and Tropical Medicine, Medical Microbiology and Pharmacology, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, HI, USA
[b] Département Systématique et Evolution, Muséum National d'Histoire Naturelle, UMR CNRS 7205, Paris, France

ARTICLE INFO

Article history:
Received 13 June 2013
Received in revised form 11 August 2013
Accepted 19 August 2013
Available online 27 August 2013

Keywords:
Hantavirus
Crocidura
Shrew
Phylogeny
Guinea

ABSTRACT

Elucidation of the molecular phylogeny of shrew-borne hantaviruses in sub-Saharan Africa has been hampered by the lack of full-length viral genomes. In this report, we present the complete genome analysis of a newfound hantavirus, designated Bowé virus, detected in ethanol-fixed intercostal muscle of a Doucet's musk shrew (*Crocidura douceti*), captured in southwestern Guinea in February 2012. Full-length amino acid sequence comparison of the S-, M- and L-segment gene products revealed that Bowé virus differed by 24.1-53.4%, 17.0-59.9% and 14.6-39.7%, respectively, from all other representative rodent-, shrew- and mole-borne hantaviruses. Phylogenetic analysis, using maximum-likelihood and Bayesian methods, under the GTR + I + Γ model of evolution, showed that Bowé virus shared a common ancestry with Tanganya virus, a hantavirus detected in the Therese's shrew (*Crocidura theresae*) in Guinea. Whole genome analysis of many more hantaviruses from sub-Saharan Africa are needed to better clarify how the radiation of African shrews might have contributed to the phylogeography of hantaviruses.

1. Introduction

Hantaviruses (genus *Hantavirus*), the etiologic agents of hemorrhagic fever with renal syndrome (HFRS) and hantavirus cardiopulmonary syndrome (HCPS), possess a single-stranded, negative-sense RNA genome, consisting of three segments, designated S, M and L, which encode the nucleocapsid protein, envelope glycoproteins and RNA-dependent RNA polymerase, respectively (Bi et al., 2008; Schmaljohn and Hjelle, 1997). Unlike other members of the Bunyaviridae family, hantaviruses have no arthropod host and have been long known to be harbored by rodents of multiple species in Europe, Asia and the Americas (Jonsson et al., 2010). Until recently, no hantavirus had been reported from Africa. This impasse was broken with the discovery of Sangassou virus (SANGV) in the African wood mouse (*Hylomyscus simus*) (Klempa et al., 2006) and Tanganya virus (TGNV) in the Therese's shrew (*Crocidura theresae*), captured in Guinea (Klempa et al., 2007). More recently, another genetically distinct hantavirus, Azagny virus (AZGV), was detected in the West African pygmy shrew (*Crocidura obscurior*) in neighboring Côte d'Ivoire (Kang et al., 2011). However, the molecular phylogeny of crocidurine shrew-borne hantaviruses from Africa has been hampered because the genomes of TGNV and AZGV have not been fully sequenced.

In an attempt to obtain the complete genomes of TGNV and AZGV, ethanol-fixed, archival intercostal muscle tissues, collected from 39 West African pygmy shrews and 22 Therese's shrews, as well as 11 Doucet's musk shrew (*Crocidura douceti*) and 11 large-headed shrews (*Crocidura grandiceps*), which were captured as part of an ecological survey of small mammals in Bowé, Gbaolé, Hermakono and Zogota in Guinea (Fig. 1A), during May 2011–February 2012, were analyzed for hantavirus RNA by reverse transcription polymerase chain reaction (RT-PCR), using previously described methods (Arai et al., 2008, 2012; Kang et al., 2009, 2011). In this report, we present the molecular phylogeny and complete genome sequence of a previously unrecognized hantavirus, designated Bowé virus (BOWV), harbored by the Doucet's musk shrew.

2. Materials and methods

Total RNA was extracted, using the PureLink Micro-to-Midi total RNA purification kit (Invitrogen, San Diego, CA, USA), from shrew tissues, and cDNA was prepared using the SuperScript III First-Strand Synthesis System (Invitrogen) and random hexamers and/or an oligonucleotide primer (OSM55: 5′-TAGTAGTAGACTCC-3′)

* Corresponding author. Fax: +1 808 6921976.
E-mail address: ryanagih@hawaii.edu (R. Yanagihara).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.08.016

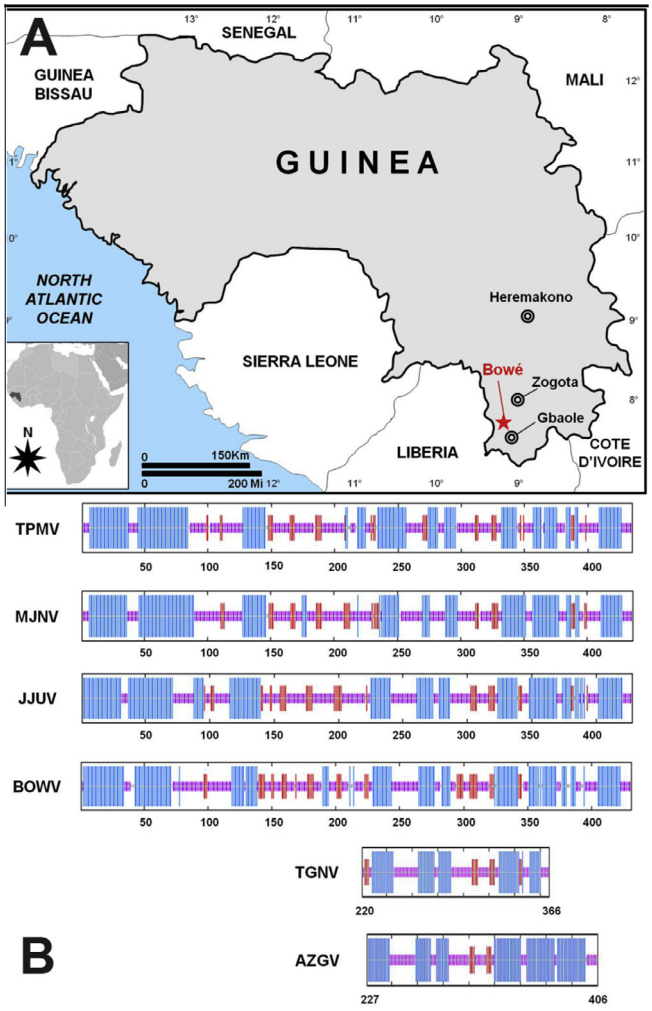

Fig. 1. (A) Map of Guinea, showing the locations of the four trap sites. A single Doucet's musk shrew captured in Bowé was found to be infected with a previously unrecognized hantavirus. (B) Schematic comparison of consensus secondary structures of the full-length N protein of TPMV, MJNV, JJUV and BOWV, with the partial N protein of AZGV and TGNV. N protein structures for the crocidurine shrew-borne hantaviruses from Africa and Asia were predicted using several methods, including HNN, DSC, PHD, PREDATOR and MLRC, available at the NPS@structure server (Combet et al., 2000). Alpha helices are represented by purple bars, beta-strand by yellow-green bars, and random coil and unclassified structure by red and gray bars, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

designed from the conserved 5′-end of the S, M and L segments of hantaviruses. PCR was performed as described previously, with each 20-μL reaction containing 250 μM dNTP, 2 mM $MgCl_2$, 1 U of AmpliTaq polymerase (Roche, Basel, Switzerland) and 0.25 μM of oligonucleotide primers, with modified cycling conditions. Innumerable trial-and-error testing of primers was necessary before ultimately obtaining the complete genome sequence of a newfound hantavirus (Table 1). Initial denaturation at 94 °C for 5 min was followed by two cycles each of denaturation at 94 °C for 40 s, two-degree step-down annealing from 48 °C to 38 °C for 40 s, and elongation at 72 °C for 1 min, then 32 cycles of denaturation at 94 °C for 40 s, annealing at 42 °C for 40 s, and elongation at 72 °C for 1 min, in a GeneAmp PCR 9700 thermal cycler (Perkin-Elmer, Waltham, MA). Amplicons were purified using the QIAQuick Gel Extraction Kit (Qiagen, Hilden, Germany), and DNA sequencing was performed using an ABI Prism 377XL Genetic Analyzer (Applied Biosystems, Foster City, CA).

In addition to traditional Sanger sequencing, BOWV RT-PCR products were sequenced using Ion Torrent PGM (Life Technologies, Inc., Carlsbad, CA) to validate the complete BOWV genome. Amplicons were size selected on 0.8% agarose gels, then purified with QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The purified DNA was quantified with Quant-iT™ Technology (Life Technologies, Inc.) using Quant-iT™ dsDNA Broad-Range Assay Kit. Libraries were then constructed using NEBNext Fast DNA Fragmentation & Library Prep Set for Ion Torrent™ (New England Biolabs, Inc., Ipswich, MA), and the Agilent Bioanalyzer™ with the Agilent High Sensitivity DNA Kit (Agilent Technologies, Inc., Santa Clara, CA) was used to analyze library sizes and molar concentrations. Template preparation and sequencing reactions were performed with Ion PGM™ 200 Xpress™ Template Kit and Ion PGM™ 200 Sequencing Kit (Life Technologies, Inc.), according to the manufacturer's instructions.

Ion Torrent base sequence quality below 20 were trimmed and reads shorter than 30 bases were discarded before further analysis. Sequence alignment was achieved using Bowtie 2–2.0.0 (Langmead and Salzberg, 2012), and Velvet 1.2.07 *de novo* assembler (Zerbino and Birney, 2008), with varying stringencies, was used to generate sequence contigs. Sequence comparison analysis for Ion Torrent sequencing was performed using BLAST (National Center for Biotechnology Information, Bethesda, MD), and ClustalW (DNASTAR, Inc., Madison, WI) (Thompson et al., 1994) and transAlign (Bininda-Emonds, 2005) were used for Sanger sequencing.

Phylogenetic trees were generated by maximum likelihood and Bayesian methods, implemented in PAUP* (Phylogenetic Analysis Using Parsimony, 4.0b10) (Swofford, 2003), RAxML Blackbox webserver (Stamatakis et al., 2008) and MrBayes 3.1 (Ronquist and Huelsenbeck, 2003), under the best-fit GTR + I + Γ model of evolution selected by hierarchical likelihood-ratio test in MrModeltest v2.3 (Posada and Crandall, 1998) and jModelTest version 0.1 (Posada, 2008). Two replicate Bayesian Metropolis–Hastings Markov Chain Monte Carlo runs, each consisting of six chains of 10 million generations sampled every 100 generations with a burn-in of 25,000 (25%), resulted in 150,000 trees overall. Each genomic segment (S, M and L) was treated separately in phylogenetic analyses. Topologies were evaluated by bootstrap analysis of 1000 iterations, and posterior node probabilities were based on 2 million generations and estimated sample sizes over 100 (implemented in MrBayes).

3. Results and discussion

Disappointingly, TGNV and AZGV were not detected in any of the Therese's shrews and West African pygmy shrews, respectively, despite using TGNV- and AZGV-specific primers. As such, we were unable to definitively establish the host designation of TGNV and AZGV. The failure to detect TGNV and AZGV may be attributed to the ethanol fixation and the less than optimal tissue (intercostal muscle), as well as the small sample size and the likely low virus titers. Nevertheless, BOWV was identified in a Doucet's musk shrew, captured on February 20, 2012 in a regenerating forested area in Bowé (8°5'60"N, 8°49'60"W), located in the Prefecture de Nzerekore region of Guinea (Fig. 1A). The taxonomic identity of the BOWV-infected Doucet's musk shrew was confirmed by mitochondrial DNA sequence analysis of the 1140-nucleotide full-length cytochrome *b* gene (GenBank accession No. KC684929).

Table 1
Oligonucleotide primers used to amplify and sequence BOWV.

Segment	Primer name	Position	Primer sequence (5′–3′)
S	OSM55	1	TAG TAG TAG ACT CC
	Shrew-S764R	680–697	CCA TNA CWG GRC TNA TCA
	Han-S604F	604–624	GCH GAD GAR HTN ACA CCN GG
	SO-S1220R	1216–1235	TCR TCA CCN AGR TGG AAR TG
	Shrew-S1174F	1108–1127	TCM TCA TTY TAY CAA TCA TA
	Shrew-end-EcoRI	1731	CTC GAA TTC TAG TAG T
M	OSM55	1	TAG TAG TAG ACT CC
	BOWV-M1594R	1563–1581	CTA CAA GTA TGC ACC ATG T
	HTN-M1520F	1537–1555	TTG GNT GGN TNY TAA THC C
	HTN-M2355R	2352–2374	CCW GGR CAA TCH VGA GGR TTA CA
	BOWV-M2318F	2295–2312	TGG CAA GCA GCA GCA TGT
	BOWV-M2856R	2806–2825	ATC TAT AGT AGC CAT AAG CT
	Shrew-end-EcoRI	3597	CTC GAA TTC TAG TAG T
L	OSM55	1	TAG TAG TAG ACT CC
	BOWV-L1981R	1913–1931	CAG ATG GTA TTA AAT ATC T
	Han-L1900F	1883–1905	ATG AAR NTN TGT GCN ATN TTT GA
	Han-L2970R	2948–2970	CCN GGN GAC CAY TTN GTD GCA TC
	Han-L-F1	2933–2952	ATG TAY GTB AGT GCW GAT GC
	Han-L5239R	5162–5181	TGN AYR CAR TAW GCA TCA TA
	BOWV-L4954F	4954–4978	AGC CAA TGA TGC ACC ATA TAT AGA C
	Shrew-3'endR	6551	TAG TAG TAK RCT CCY TRA A

The poor-quality fragmented RNA, extracted from the ethanol-fixed intercostal muscle, thwarted virus-specific RNA enrichment or host-sequence depletion efforts for RNA sequencing. Therefore, we performed deep sequencing of viral RT-PCR amplicons. Sequencing of the BOWV genome by Ion Torrent generated 2,777,993 reads with a mean read length of 113 bases. The sequence reads were trimmed and filtered for quality before being aligned to Sanger sequencing accomplished sequences and *de novo* assembled. Mapping sequence reads revealed 1,499,052 hits (55% of the quality reads) with mean coverage alignment of 13,319×; 244,726 reads (16.3%) mapped to the S segment, 501,888 reads (33.5%) to the M segment, and 752,438 reads (50.2%) to the L segment. Ultimately, the consensus sequences were accomplished from the alignment, which covered 73.4% of the BOWV genome. Comparison of the assembled contigs, consensus sequences by both Sanger and next-generation sequencing revealed more than 99% similarity for each of the segments. GenBank accession numbers for the S, M and L segments of BOWV were KC631782, KC631783 and KC631784, respectively.

The full-length 1731-nucleotide S-genomic segment of BOWV contained a single open reading frame (ORF), encoding a 431-amino acid nucleocapsid (N) protein (nucleotide positions, 28–1,320), and 3′- and 5′-noncoding regions (NCR) of 27- and 408-nucleotides, respectively. The hypothetical NSs ORF, typically found in hantaviruses hosted by arvicolid, neotomine and sigmodontine rodents, was absent. Employing prediction software available in the NPS@structure server (Combet et al., 2000), the BOWV N protein secondary structure generally resembled that of other rodent- and soricomorph-borne hantaviruses, comprising two α-helical domains packed against a central β-pleated sheet. Direct comparison of the N protein secondary structure demonstrated that BOWV most closely resembled AZGV, TGNV and Jeju virus (JJUV), and differed from Thottapalayam virus (TPMV) (Carey et al., 1971) and Imjin virus (MJNV) (Song et al., 2009), two crocidurine shrew-borne hantaviruses from Asia (Fig. 1B).

The complete M-genomic segment of BOWV was 3597 nucleotides, with a predicted glycoprotein of 1145 amino acids, starting at nucleotide position 36, and a 124-nucleotide 5′-NCR. Like other hantaviruses, the BOWV glycoprotein precursor had the highly conserved WAASA amino-acid motif (amino acid positions 652–656). Glycosylation sites, as predicted using NetNlyc 1.0 and Predictprotein (Gavel and von Heijne, 1990), showed five potential N-linked glycosylation sites (four in Gn at amino acid positions 142, 355, 407 and 526; and one in Gc at position 936).

The full-length 6,551-nucleotide L-genomic segment of BOWV encoded a predicted RNA-dependent RNA polymerase of 2158 amino acids. The six major conserved motifs (designated as premotif A and motifs A, B, C, D and E) for RNA virus polymerases, which are shared by all hantaviruses, were also found in BOWV. Premotif A had a conserved lysine and two arginine residues. Motifs A, B and D had conserved aspartate, glycine and lysine, respectively. In motif C, there were two conserved aspartic acid residues. The XDD motif, essential for catalytic activity, and motif E, containing the E(F/Y)XS site, were also present.

Pairwise alignment and comparison of the S-segment coding sequences of BOWV and other non-African soricomorph-borne hantaviruses showed moderately low sequence similarities, ranging from 48.2-71.3% and 46.6-75.9% at the nucleotide and amino acid levels, respectively (Table 2). Amino acid sequence identity of the RNA-dependent RNA polymerase was highest (83.1% and 85.4%) between BOWV and AZGV and TGNV, two other African crocidurine shrew-borne hantaviruses (Table 2). This overall high sequence similarity was consistent with the presumed functional constraints on the RNA-dependent RNA polymerase.

The disparate phylogenetic positions of BOWV in trees based on the S- and on the M- and L-segment sequences could not be attributed to reassortment or recombination. Multiple recombination-detection methods, including GENECONV, Bootscan, Chimaera, 3SEQ, RDP, SiScan, MaxChi and HyPhy Single Recombinant Breakpoint (Pond et al., 2005), failed to disclose any evidence of genetic recombination in the BOWV genome.

Unrooted phylogenetic trees, based on the nucleotide sequences, showed four distinct, well-supported clades, comprising hantaviruses harbored by rodents of the Muridae family; rodents of the Cricetidae family; soricomorphs of the Soricidae and Talpidae families; and two shrew species (Crocidurinae subfamily)

Table 2
Nucleotide and amino acid sequence identity (%) between BOWV strain VN1512 and representative rodent-, shrew- and mole-borne hantaviruses.

Virus strain	S segment		M segment		L segment	
	1293 nt	431 aa	3435 nt	1145 aa	6474 nt	2158 aa
HTNV 76-118	66.4	64.3	62.9	58.1	69.7	73.4
SEOV 80-39	63.9	63.6	61.2	56.9	69.1	72.3
SANGV SA14	65.3	63.2	60.7	56.7	68.4	72.0
SOOV SOO-1	64.0	63.6	63.3	58.2	70.2	73.2
DOBV Greece	65.0	62.2	62.6	56.8	69.9	73.5
PHV PH-1	63.3	58.0	59.2	50.7	66.4	65.4
PUUV Sotkamo	63.3	61.5	60.5	51.1	67.1	66.6
TULV 5302v	63.2	60.4	59.2	51.8	66.5	65.3
ANDV Chile9717869	65.0	61.4	59.6	51.7	67.2	66.5
SNV NMH10	61.4	59.1	59.3	51.8	66.2	66.7
CBNV CBN-3	69.0	74.5	66.5	66.2	72.2	78.2
ARRV MSB73418	67.3	65.6	–	–	71.1	73.9
JMSV MSB144475	69.3	71.8	70.2	71.7	72.1	78.8
SWSV mp70	66.9	67.1	71.5	73.5	68.8	74.2
QDLV YN05-284	67.1	67.6	68.3	68.9	74.8	79.9
KKMV MSB148794	65.7	67.6	70.7	68.6	71.3	77.8
TGNV Tan826	72.2	73.5	–	–	76.2	85.4
AZGV KBM15	65.0	65.6	73.1	83.0	74.2	83.1
JJUV 10-11	71.3	75.9	69.7	71.7	73.6	80.6
MJNV Cl05-11	49.1	47.6	41.5	41.1	64.4	60.3
TPMV VRC66412	48.2	46.6	40.3	40.1	63.5	60.5
RKPV MSB57412	64.3	61.2	62.1	51.3	67.8	65.8
OXBV Ng1453	68.1	70.3	64.7	61.6	70.4	75.7
ASAV N10	68.5	69.6	67.9	64.4	72.1	77.0
NVAV MSB95703	48.6	48.8	55.1	42.4	64.0	60.4

Abbreviations: ANDV, Andes virus; ARRV, Ash River virus; ASAV, Asama virus; AZGV, Azagny virus; CBNV, Cao Bang virus; DOBV, Dobrava virus; HTNV, Hantaan virus; JMSV, Jemez Spring virus; JJUV, Jeju virus; KKMV, Kenkeme virus; MJNV, Imjin virus; NVAV, Nova virus; OXBV, Oxbow virus; PHV, Prospect Hill virus; PUUV, Puumala virus; QDLV, Qiandao Lake virus; RKPV, Rockport virus; SANGV, Sangassou virus; SEOV, Seoul virus; SNV, Sin Nombre virus; SOOV, Soochong virus; SWSV, Seewis virus; TGNV, Tanganya virus; TPMV, Thottapalayam virus; TULV, Tula virus. nt, nucleotides; aa, amino acids.

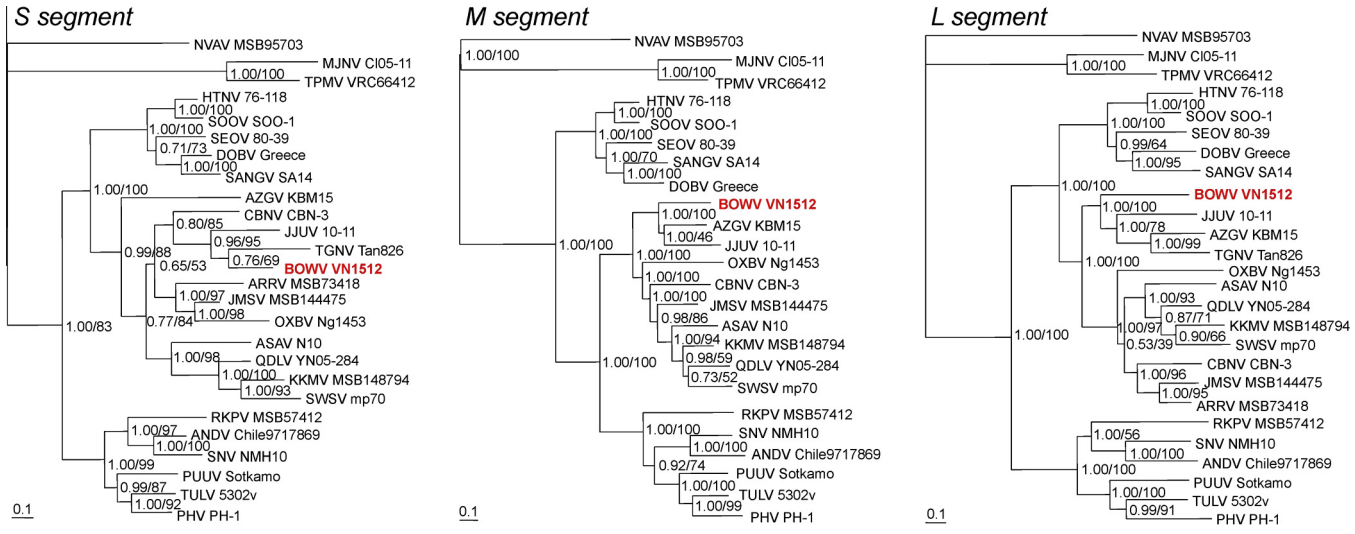

Fig. 2. Phylogenetic trees were generated by the maximum-likelihood and Bayesian methods, using the GTR + I + Γ model of evolution, based on the alignment of the S-, M- and L-segment sequences of BOWV strain VN1512. Since tree topologies were very similar using RAxML and MrBayes, the trees generated by MrBayes were displayed. The phylogenetic positions of BOWV are shown in relationship to crocidurine shrew-borne hantaviruses, including Tanganya virus (TGNV Tan826: EF050455, EF050454), Azagny virus (AZGV KBM15: JF276226, JF276227, JF276228), Jeju virus (JJUV 10-11: HQ834695, HQ834696, HQ834697), Thottapalayam virus (TPMV VRC66412: AY526097, EU001329, EU001330) and Imjin virus (MJNV Cl05-11: EF641804, EF641798, EF641806). Soricine shrew-borne hantaviruses included Cao Bang virus (CBNV CBN-3: EF543524, EF543526, EF543525), Ash River virus (ARRV MSB73418: EF650086, EF619961), Jemez Springs virus (JMSV MSB144475: FJ593499, FJ593500, FJ593501), Kenkeme virus (KKMV MSB148794: GQ306148, GQ306149, GQ306150), Qiandao Lake virus (QDLV YN05-284: GU566023, GU566022, GU566021) and Seewis virus (SWSV mp70: EF636024, EF636025, EF636026). Also shown are mole-borne hantaviruses, including Asama virus (ASAV N10: EU929072, EU929075, EU929078), Nova virus (NVAV MSB95703: FJ539168, HQ840957, FJ593498), Oxbow virus (OXBV Ng1453: FJ539166, FJ539167, FJ593497) and Rockport virus (RKPV MSB57412: HM015223, HM015219, HM015221). Rodent-borne hantaviruses included Hantaan virus (HTNV 76–118: NC_005218, Y00386, NC_005222), Soochong virus (SOOV SOO-1: AY675349, AY675353, DQ562292), Dobrava virus (DOBV Greece: NC_005233, NC_005234, NC_005235), Seoul virus (SEOV 80-39: NC_005236, NC_005237, NC_005238), Sangassou virus (SANGV SA14: JQ082300, JQ082301, JQ082302), Tula virus (TULV M5302v: NC_005227, NC_005228, NC_005226), Puumala virus (PUUV Sotkamo: NC_005224, NC_005223, NC_005225), Prospect Hill virus (PHV PH-1: Z49098, X55129, EF646763), Andes virus (ANDV Chile9717869: NC_003466, NC_003467, NC_003468) and Sin Nombre virus (SNV NMH10: NC_005216, NC_005215, NC_005217). The numbers at each node are posterior node probabilities based on 150,000 trees (left) and bootstrap values of 1,000 replicates executed on the RAxML BlackBox web server (right), respectively. The scale bars indicate nucleotide substitutions per site.

and one talpid mole species (Talpinae subfamily), which represent the most divergent hantaviruses found to date. The phylogenetic trees obtained for the M and L segments showed that the three African crocidurine shrew-borne hantaviruses (BOWV, TGNV and AZGV) clustered with JJUV (Arai et al., 2012), hosted by the Asian lesser white-toothed shrew (*Crocidura shantungensis*) in Korea (Fig. 2). This crocidurine shrew-borne hantavirus lineage was very well supported, and BOWV always occupied a basal position within this cluster. By contrast, as discussed previously, the divergent and basal position of MJNV and TPMV in the phylogeny of hantaviruses harbored by crocidurine shrews suggested that it may have arisen from a host-switching event (Arai et al., 2012).

Visceral tissues, such as lung and spleen, were unavailable from the hantavirus-infected Doucet's musk shrew. However, the detection of hantavirus RNA in intercostal muscle was consistent with widespread persistent infection and strongly suggests a genuine hantavirus-reservoir host relationship. Moreover, although not ideal and far inferior to frozen tissues or tissues preserved in RNAlater®, ethanol-fixed archival tissues from small mammals hold promise in virus-discovery efforts, as has been demonstrated recently with the identification of a highly divergent hantavirus in ethanol-fixed liver tissue collected from banana pipistrelles (*Neoromicia nanus*) captured in Côte d'Ivoire (Sumibcay et al., 2012). That is, ethanol-fixed tissues, originally collected for other purposes, serve as precious archival research resources in microbiology, evolutionary biology and host genetics.

The Doucet's musk shrew is a semi-arboreal species, which inhabits the forest-savanna border of Guinea and Côte d'Ivoire. Sparse populations of this species have been recorded in Adiopodoumé and the Taï National Park in Côte d'Ivoire, and in Sérédou and at the foot of Mount Nimba in upper Guinea. Of the 11 Doucet's musk shrews from Guinea in the present study, six were trapped in Gbaolé, three in Zogota and one each in Hermakono and Bowé. Although not currently listed as a threatened or endangered species by the International Union for Conservation of Nature and Natural Resources, general deforestation is a presumed threat. More information is needed about the natural history and ecology of this little-known African crocidurine shrew species.

Genome-wide analyses of many more shrew-borne hantaviruses are necessary to ascertain their evolutionary origin and phylogeography. Viewed within the context that newly recognized hantaviruses detected in African shrews greatly outnumber those reported in rodents (Klempa et al., 2006; Meheretu et al., 2012), the recent identification of highly divergent lineages of hantaviruses in insectivorous bats in Côte d'Ivoire (Sumibcay et al., 2012) and Sierra Leone (Weiss et al., 2012), as well as in China (Guo et al., 2013) and Vietnam (Arai et al., 2013), further supports the emerging concept that rodents were not the primordial hosts of ancestral hantaviruses.

Acknowledgments

This work was supported in part by U.S. Public Health Service grants R01AI075057 from the National Institute of Allergy and Infectious Diseases and P20GM103516 from the National Institute of General Medical Sciences, National Institutes of Health. We also thank the Service de Systématique Moléculaire of the Muséum National d'Histoire Naturelle (CNRS - UMS 2700, Paris, France). Trapping of shrews was approved by the Ministère de l'environnement et du développement durable, République de Guinée (permit numbers 0003394, 0003329 and 0003374).

References

Arai, S., Ohdachi, S.D., Asakawa, M., Kang, H.J., Mocz, G., Arikawa, J., Okabe, N., Yanagihara, R., 2008. Molecular phylogeny of a newfound hantavirus in the Japanese shrew mole (*Urotrichus talpoides*). Proc. Natl. Acad. Sci. U.S.A. 105, 16296–16301.

Arai, S., Gu, S.H., Baek, L.J., Tabara, K., Bennett, S.N., Oh, H.-S., Takada, N., Kang, H.J., Tanaka-Taya, K., Morikawa, S., Okabe, N., Yanagihara, R., Song, J.-W., 2012. Divergent ancestral lineages of newfound hantaviruses harbored by phylogenetically related crocidurine shrew species in Korea. Virology 424, 99–105.

Arai, S., Nguyen, S.T., Boldgiv, B., Fukui, D., Araki, K., Dang, C.N., Ohdachi, S.D., Nguyen, N.X., Pham, T.D., Boldbaatar, B., Satoh, H., Yoshikawa, Y., Morikawa, S., Tanaka-Taya, K., Yanagihara, R., Oishi, K., 2013. Novel bat-borne hantavirus. Vietnam. Emerg. Infect. Dis. 19, 1159–1161.

Bi, Z., Formenty, P.B., Roth, C.E., 2008. Hantavirus infection: a review and global update. J. Infect. Dev. Ctries. 2, 3–23.

Bininda-Emonds, O.R., 2005. TransAlign: using amino acids to facilitate the multiple alignment of protein-coding DNA sequences. BMC Bioinf. 6, 156.

Carey, D.E., Reuben, R., Panicker, K.N., Shope, R.E., Myers, R.M., 1971. Thottapalayam virus: a presumptive arbovirus isolated from a shrew in India. Indian J. Med. Res. 59, 1758–1760.

Combet, C., Blanchet, C., Geourjon, C., Deléage, G., 2000. NPS@: network protein sequence analysis. Trends Biochem. Sci. 25, 147–150.

Gavel, Y., von Heijne, G., 1990. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. Protein Eng. 3, 433–442.

Guo, W.P., Lin, X.D., Wang, W., Tian, J.H., Cong, M.L., Zhang, H.L., Wang, M.R., Zhou, R.H., Wang, J.B., Li, M.H., Xu, J., Holmes, E.C., Zhang, Y.Z., 2013. Phylogeny and origins of hantaviruses harbored by bats, insectivores, and rodents. PLoS Pathog. 9, e1003159.

Jonsson, C.B., Figueiredo, L.T., Vapalahti, O., 2010. A global perspective on hantavirus ecology, epidemiology, and disease. Clin. Microbiol. Rev. 23, 412–441.

Kang, H.J., Bennett, S.N., Sumibcay, L., Arai, S., Hope, A.G., Mocz, G., Song, J.-W., Cook, J.A., Yanagihara, R., 2009. Evolutionary insights from a genetically divergent hantavirus harbored by the European common mole (*Talpa europaea*). PLoS One 4, e6149.

Kang, H.J., Kadjo, B., Dubey, S., Jacquet, F., Yanagihara, R., 2011. Molecular evolution of azagny virus, a newfound hantavirus in the West African pygmy shrew (*Crocidura obscurior*) in Côte d'Ivoire. Virol. J. 8, 373.

Klempa, B., Fichet-Calvet, E., Lecompte, E., Auste, B., Aniskin, V., Meisel, H., Denys, C., Koivogui, L., ter Meulen, J., Krüger, D.H., 2006. Hantavirus in African wood mouse. Guinea. Emerg. Infect. Dis. 12, 838–840.

Klempa, B., Fichet-Calvet, E., Lecompte, E., Auste, B., Aniskin, V., Meisel, H., Barriere, P., Koivogui, L., ter Meulen, J., Krüger, D.H., 2007. Novel hantavirus sequences in shrew. Guinea. Emerg. Infect. Dis. 13, 520–522.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Meheretu, Y., Cížková, D., Těšíková, J., Welegerima, K., Tomas, Z., Kidane, D., Girmay, K., Schmidt-Chanasit, J., Bryja, J., Günther, S., Bryjová, A., Leirs, H., Goüy de Bellocq, J., 2012. High diversity of RNA viruses in rodents. Ethiopia. Emerg. Infect. Dis. 18, 2047–2050.

Pond, S.L., Frost, S.D.W., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679.

Posada, D., 2008. JModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Schmaljohn, C., Hjelle, B., 1997. Hantaviruses: a global disease problem. Emerg. Infect. Dis. 3, 95–104.

Song, J.-W., Kang, H.J., Gu, S.H., Moon, S.S., Bennett, S.N., Song, K.J., Baek, L.J., Kim, H.C., O'Guinn, M.L., Chong, S.T., Klein, T.A., Yanagihara, R., 2009. Characterization of Imjin virus, a newly isolated hantavirus from the ussuri white-toothed shrew (*Crocidura lasiura*). J. Virol. 83, 6184–6191.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web-servers. Syst. Biol. 75, 758–771.

Sumibcay, L., Kadjo, B., Gu, S.H., Kang, H.J., Lim, B.K., Cook, J.A., Song, J.-W., Yanagihara, R., 2012. Divergent lineage of a novel hantavirus in the banana pipistrelle (*Neoromicia nanus*) in Côte d'Ivoire. Virol. J. 9, 34.

Swofford, D.L., 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence

weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Weiss, S., Witkowski, P.T., Auste, B., Nowak, K., Weber, N., Fahr, J., Mombouli, J.V., Wolfe, N.D., Drexler, J.F., Drosten, C., Klempa, B., Leendertz, F.H., Krüger, D.H., 2012. Hantavirus in bat, Sierra Leone. Emerg. Infect. Dis. 18, 159–161.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829.